

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Hillman, Jennifer L.
 Shah, Purvi
 Corley, Neil C.

(ii) TITLE OF THE INVENTION: HUMAN PEROXISOMAL THIOESTERASE

(iii) NUMBER OF SEQUENCES: 4

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
 (B) STREET: 3174 Porter Drive
 (C) CITY: Palo Alto
 (D) STATE: CA
 (E) COUNTRY: USA
 (F) ZIP: 94304

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette
 (B) COMPUTER: IBM Compatible
 (C) OPERATING SYSTEM: DOS
 (D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: To Be Assigned
 (B) FILING DATE: Filed Herewith

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER:
 (B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Billings, Lucy J.
 (B) REGISTRATION NUMBER: 36,749
 (C) REFERENCE/DOCKET NUMBER: PF-0293 US

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 415-855-0555
 (B) TELEFAX: 415-845-4166

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 311 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: BRAINOT09
 (B) CLONE: 2150905

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Met Gly Arg Ala Val Ala Thr Ala Ala Leu Pro Pro Gly Asp Leu Arg
 1 5 10 15
 Ser Val Leu Val Thr Thr Val Leu Asn Leu Glu Pro Leu Asp Glu Asp
 20 25 30
 Leu Phe Arg Gly Arg His Tyr Trp Val Pro Ala Lys Arg Leu Phe Gly
 35 40 45

Gly Gln Ile Val Gly Gln Ala Leu Val Ala Ala Ala Lys Ser Val Ser
 50 55 60
 Glu Asp Val His Val His Ser Leu His Cys Tyr Phe Val Arg Ala Gly
 65 70 75 80
 Asp Pro Lys Leu Pro Val Leu Tyr Gln Val Glu Arg Thr Arg Thr Gly
 85 90 95
 Ser Ser Phe Ser Val Arg Ser Val Lys Ala Val Gln His Gly Lys Pro
 100 105 110
 Ile Phe Ile Cys Gln Ala Ser Phe Gln Gln Ala Gln Pro Ser Pro Met
 115 120 125
 Gln His Gln Phe Ser Met Pro Thr Val Pro Pro Pro Glu Glu Leu Leu
 130 135 140
 Asp Cys Glu Thr Leu Ile Asp Gln Tyr Leu Arg Asp Pro Asn Leu Gln
 145 150 155 160
 Lys Arg Tyr Pro Leu Ala Leu Asn Arg Ile Ala Ala Gln Glu Val Pro
 165 170 175
 Ile Glu Ile Lys Pro Val Asn Pro Ser Pro Leu Ser Gln Leu Gln Arg
 180 185 190
 Met Glu Pro Lys Gln Met Phe Trp Val Arg Ala Arg Gly Tyr Ile Gly
 195 200 205
 Glu Gly Asp Met Lys Met His Cys Cys Val Ala Ala Tyr Ile Ser Asp
 210 215 220
 Tyr Ala Phe Leu Gly Thr Ala Leu Leu Pro His Gln Trp Gln His Lys
 225 230 235 240
 Val His Phe Met Val Ser Leu Asp His Ser Met Trp Phe His Ala Pro
 245 250 255
 Phe Arg Ala Asp His Trp Met Leu Tyr Glu Cys Glu Ser Pro Trp Ala
 260 265 270
 Gly Gly Ser Arg Gly Leu Val His Gly Arg Leu Trp Arg Gln Asp Gly
 275 280 285
 Val Leu Ala Val Thr Cys Ala Gln Glu Gly Val Ile Arg Val Lys Pro
 290 295 300
 Gln Val Ser Glu Ser Lys Leu
 305 310

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1098 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: BRAINOT09
- (B) CLONE: 2150905

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CAGCATTGAA	CTAGATGTCG	TCCCCGCAGG	CCCCAGAAAGA	TGGGCAGGGC	TGTGGCAGACC	60
GCGGCGCTTC	CCCCTGGGGA	CCTCCGTAGC	GTCTTGGTCA	CGACCGTGCT	CAACCTCGAG	120
CCGCTGGACG	AGGATCTCTT	CAGAGGAAGG	CATTACTGGG	TACCGGCCAA	GAGGCTGTTT	180
GGTGGTCAGA	TCGTGGGCCA	GGCCCTGGTG	GCTGCAGCCA	AGTCTGTGAG	TGAAGACGTC	240
CACGTGCACT	CCCTGCCACTG	CTACTTTGTT	CGGGCAGGGG	ACCCGAAGCT	GCCAGTACTG	300
TACCAAGTGG	AGCGGACACG	ACAGGGTCG	AGCTTCTCGG	TGCGCTCTGT	GAAGGCCGTG	360
CAACATGGGA	AGCCCATCTT	CATCTGCCAG	GCCTCCTTCC	AGCAGGCCA	GCCAGCCCC	420
ATGCAGCACC	AGTTCTCCAT	GCCCACGTG	CCACCAACAG	AAGAGCTGCT	TGACTGTGAG	480
ACCCCTATTG	ACCACTATTG	AAGGGACCTT	AACCTCCAAA	AGAGGTACCC	ATTGGCGCTC	540
AACCGAATTG	CTGCTCAGGA	GGTCCCCATT	GAGATCAAGC	CAGTAAACCC	ATCCCCCTG	600
AGCCAGCTGC	AGAGAATGGA	GCCCAACAG	ATGTTCTGGG	TGCGAGCCCG	GGGCTATATT	660
GGCGAGGGCG	ACATGAAGAT	GCACTGCTGC	GTGGCCGCCT	ATATCTCCGA	CTATGCCTTC	720
TTGGGCACTG	CACTGCTGCC	TCACCACTGG	CAGCACAAAGG	TGCACTTCAT	GGTCTCACTG	780
GACCATTCCA	TGTGGTTCCA	CGCCCCCTTC	CGAGCTGACC	ACTGGATGCT	CTATGAATGC	840
GAGAGCCCT	GGGCCGGTGG	CTCTCGGGGG	CTGGTCCATG	GGCGGCTGTG	GCGTCAGGAT	900
GGAGTCCTAG	CTGTGACCTG	TGCCCCAGGAG	GGCGTGTATCC	GAGTGAAGCC	CCAGGTCTCA	960
GAGAGCAAGC	TGTAGCCAGA	GGTACCAAGCT	TCGCCTGGGG	CTTCAAGAAC	CTCCCCATCTA	1020

TCCCCATTCC TGAGACAGGA GTTACAGTCC CTTTGCCCC TCACATCCAA TAAAGAGACT
GATACCACTG GAAAAAAA1080
1098

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 286 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (vii) IMMEDIATE SOURCE:
- (A) LIBRARY: GenBank
 - (B) CLONE: 147932

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Ser Gln Ala Leu Lys Asn Leu Leu Thr Leu Leu Asn Leu Glu Lys
 1 5 10 15
 Ile Glu Glu Gly Leu Phe Arg Gly Gln Ser Glu Asp Leu Gly Leu Arg
 20 25 30
 Gln Val Phe Gly Gly Gln Val Val Gly Gln Ala Leu Tyr Ala Ala Lys
 35 40 45
 Glu Thr Val Pro Glu Glu Arg Leu Val His Ser Phe His Ser Tyr Phe
 50 55 60
 Leu Arg Pro Gly Asp Ser Lys Lys Pro Ile Ile Tyr Asp Val Glu Thr
 65 70 75 80
 Leu Arg Asp Gly Asn Ser Phe Ser Ala Arg Arg Val Ala Ala Ile Gln
 85 90 95
 Asn Gly Lys Pro Ile Phe Tyr Met Thr Ala Ser Phe Gln Ala Pro Glu
 100 105 110
 Ala Gly Phe Glu His Gln Lys Thr Met Pro Ser Ala Pro Ala Pro Asp
 115 120 125
 Gly Leu Pro Ser Glu Thr Gln Ile Ala Gln Ser Leu Ala His Leu Leu
 130 135 140
 Pro Pro Val Leu Lys Asp Lys Phe Ile Cys Asp Arg Pro Leu Glu Val
 145 150 155 160
 Arg Pro Val Glu Phe His Asn Pro Leu Lys Gly His Val Ala Glu Pro
 165 170 175
 His Arg Gln Val Trp Ile Arg Ala Asn Gly Ser Val Pro Asp Asp Leu
 180 185 190
 Arg Val His Gln Tyr Leu Leu Gly Tyr Ala Ser Asp Leu Asn Phe Leu
 195 200 205
 Pro Val Ala Leu Gln Pro His Gly Ile Gly Phe Leu Glu Pro Gly Ile
 210 215 220
 Gln Ile Ala Thr Ile Asp His Ser Met Trp Phe His Arg Pro Phe Asn
 225 230 235 240
 Leu Asn Glu Trp Leu Leu Tyr Ser Val Glu Ser Thr Ser Ala Ser Ser
 245 250 255
 Ala Arg Gly Phe Val Arg Gly Glu Phe Tyr Thr Gln Asp Gly Val Leu
 260 265 270
 Val Ala Ser Thr Val Gln Glu Gly Val Met Arg Asn His Asn
 275 280 285

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 349 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (vii) IMMEDIATE SOURCE:
- (A) LIBRARY: GenBank
 - (B) CLONE: 854594

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Ser Ala Ser Lys Met Ala Met Ser Asn Leu Glu Lys Ile Leu Glu
 1 5 10 15
 Leu Val Pro Leu Ser Pro Thr Ser Phe Val Thr Lys Tyr Leu Pro Ala
 20 25 30
 Ala Pro Val Gly Ser Lys Gly Thr Phe Gly Gly Thr Leu Val Ser Gln
 35 40 45
 Ser Leu Leu Ala Ser Leu His Thr Val Pro Leu Asn Phe Phe Pro Thr
 50 55 60
 Ser Leu His Ser Tyr Phe Ile Lys Gly Gly Asp Pro Arg Thr Lys Ile
 65 70 75 80
 Thr Tyr His Val Gln Asn Leu Arg Asn Gly Arg Asn Phe Ile His Lys
 85 90 95
 Gln Val Ser Ala Tyr Gln His Asp Lys Leu Ile Phe Thr Ser Met Ile
 100 105 110
 Leu Phe Ala Val Gln Arg Ser Lys Glu His Asp Ser Leu Gln His Trp
 115 120 125
 Glu Thr Ile Pro Gly Leu Gln Gly Lys Gln Pro Asp Pro His Arg Tyr
 130 135 140
 Glu Glu Ala Thr Ser Leu Phe Gln Lys Glu Val Leu Asp Pro Gln Lys
 145 150 155 160
 Leu Ser Arg Tyr Ala Ser Leu Ser Asp Arg Phe Gln Asp Ala Thr Ser
 165 170 175
 Met Ser Lys Tyr Val Asp Ala Phe Gln Tyr Gly Val Met Glu Tyr Gln
 180 185 190
 Phe Pro Lys Asp Met Phe Tyr Ser Ala Arg His Thr Asp Glu Leu Asp
 195 200 205
 Tyr Phe Val Lys Val Arg Pro Pro Ile Thr Thr Val Glu His Ala Gly
 210 215 220
 Asp Glu Ser Ser Leu His Lys His His Pro Tyr Arg Ile Pro Lys Ser
 225 230 235 240
 Ile Thr Pro Glu Asn Asp Ala Arg Tyr Asn Tyr Val Ala Phe Ala Tyr
 245 250 255
 Leu Ser Asp Ser Tyr Leu Leu Leu Thr Ile Pro Tyr Phe His Asn Leu
 260 265 270
 Pro Leu Tyr Cys His Ser Phe Ser Val Ser Leu Asp His Thr Ile Tyr
 275 280 285
 Phe His Gln Leu Pro His Val Asn Asn Trp Ile Tyr Leu Lys Ile Ser
 290 295 300
 Asn Pro Arg Ser His Trp Asp Lys His Leu Val Gln Gly Lys Tyr Phe
 305 310 315 320
 Asp Thr Gln Ser Gly Arg Ile Met Ala Ser Val Ser Gln Glu Gly Tyr
 325 330 335
 Val Val Tyr Gly Ser Glu Arg Asp Ile Arg Ala Lys Phe
 340 345